데이터분석을 위한 선형대수학

데이터를 벡터로 이해하기

김용담 강사

Contents
데이터분석을 위한 선형대수

1. 데이터분석과 선형대수

데이터 분석 예시
Q. 치즈버거와 가장 영양성분이 비슷한 메뉴를 어떻게 찾을 수 있을까?

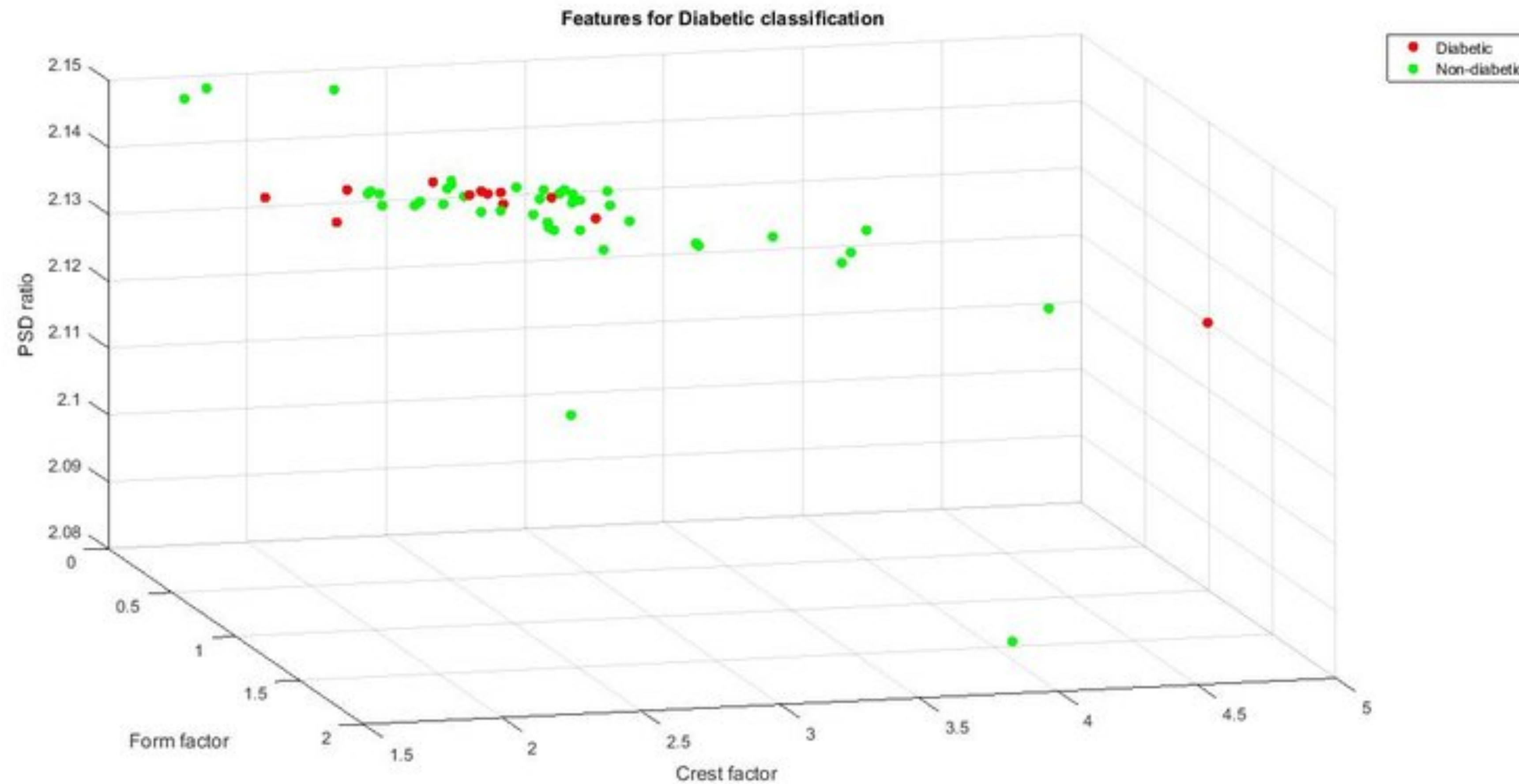| Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

데이터 분석 예시
Q. 어떤 기준으로 영양 성분의 비슷함을 말할 수 있을까?

Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

데이터 분석 예시

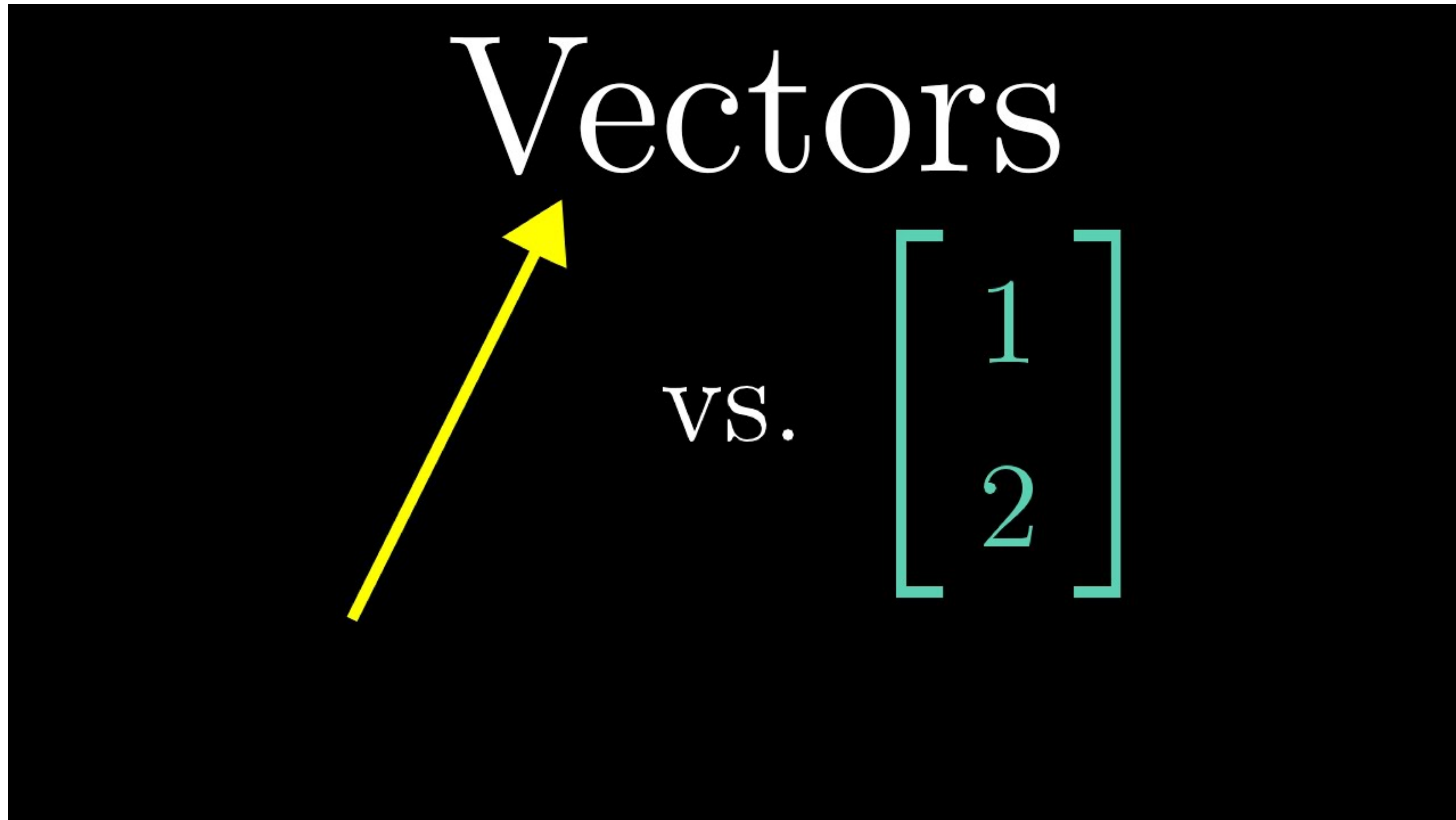**A.** 데이터는 벡터다

Features for Diabetic classification

2. 벡터란 무엇인가?

벡터의 정의
벡터 공간의 원소

벡터의 정의
여러 개의 숫자 모음

- 여러 개의 숫자를 묶어서 표현한 것

  e.g. v = $(1, 2, 3)$

- 숫자 묶음을 하나의 개념으로 정의함

  e.g. 학생 한 명 = (175, 70, 4.1)

$\cdot (1, 2, 3)$

$\boldsymbol{x} = (x_1, x_2, \cdots, x_N)$
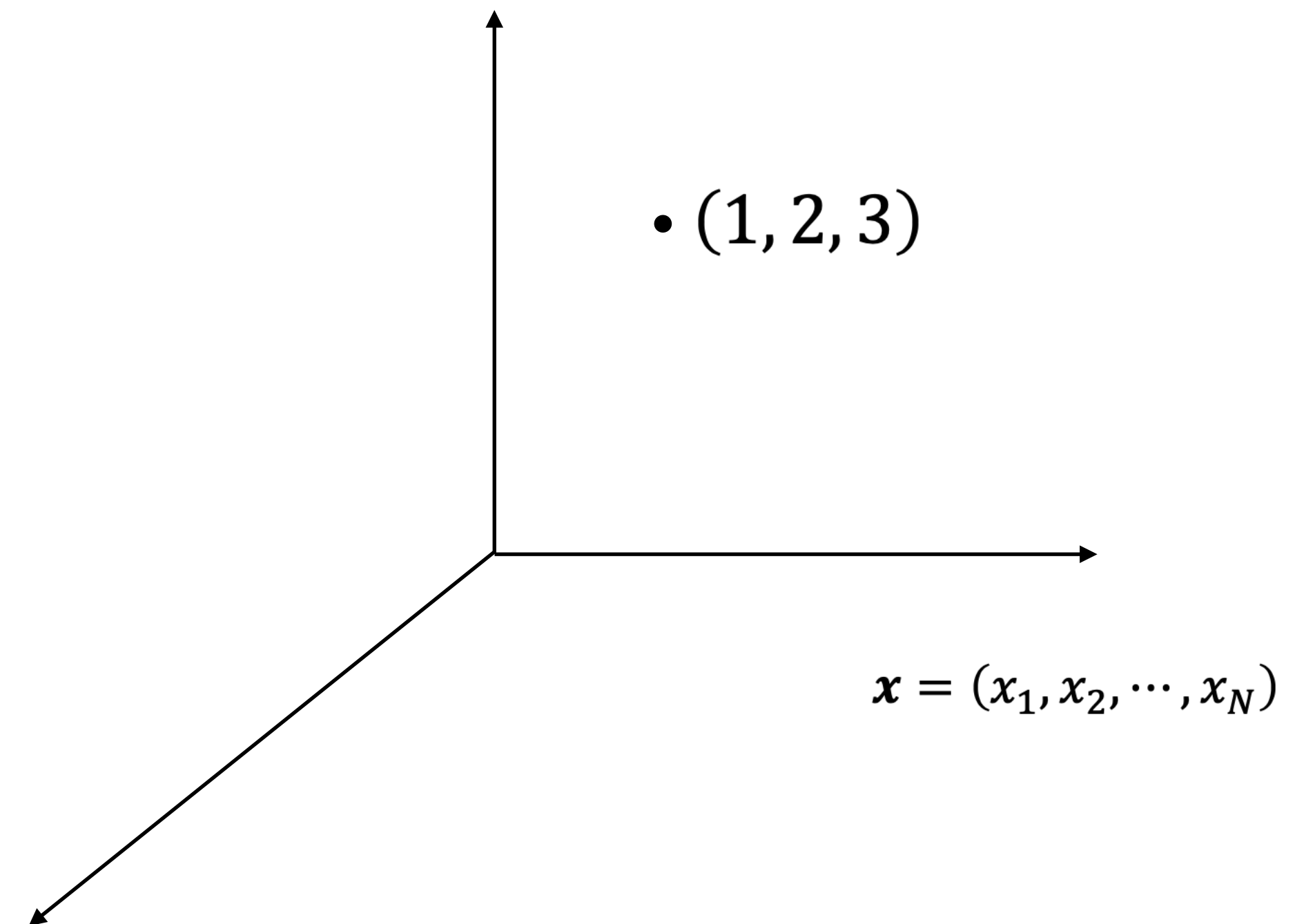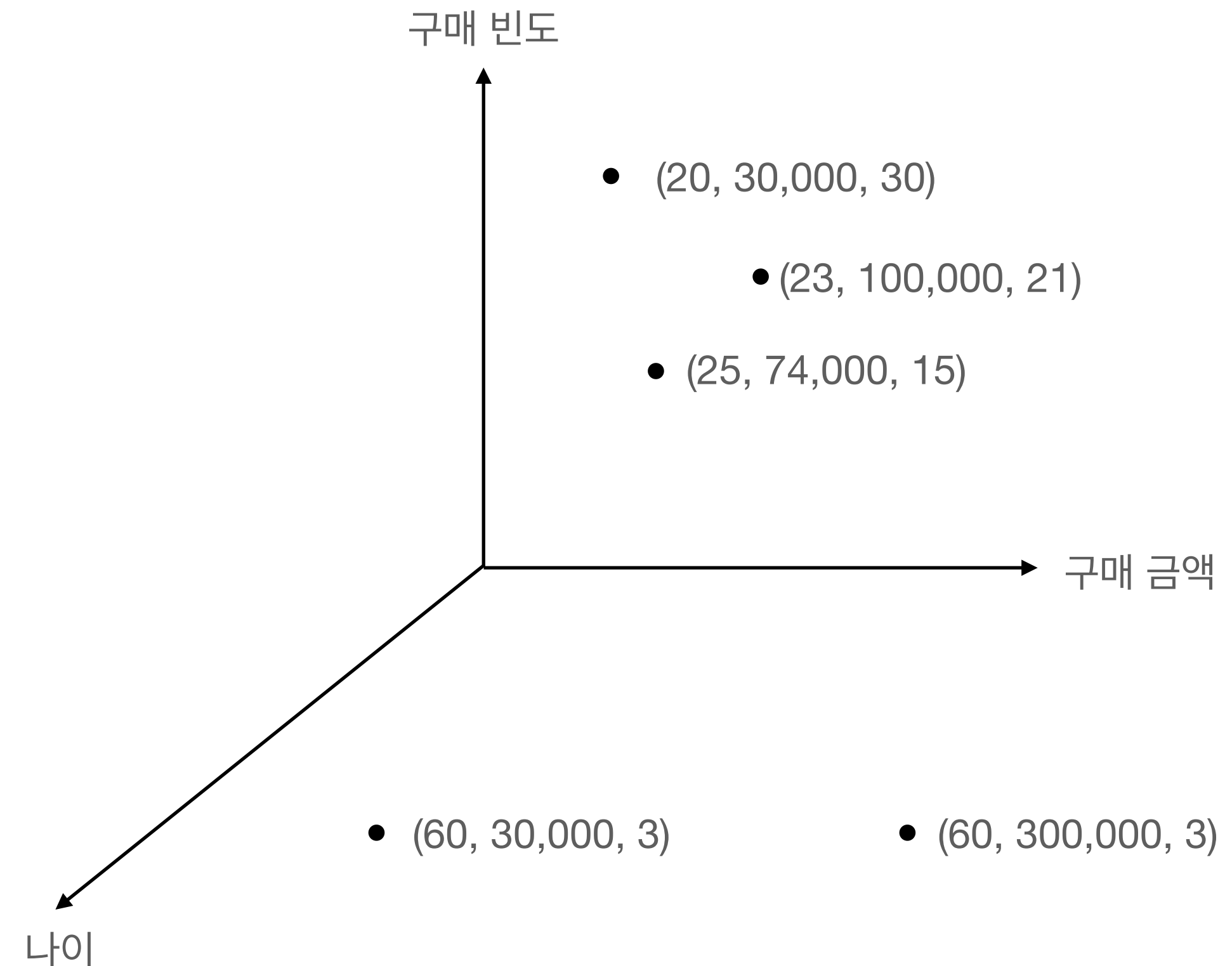
벡터와 관련된 용어들
벡터 공간(Vector Space)에 대한 이해

- 축(axis) = 벡터의 각 원소의 의미

- 차원(dimension) = 벡터의 원소 개수

- 벡터 공간이란, 축을 기준으로 어떤 의미를 가지는 숫자들의 모음을 의미한다. 몇 개의 숫자들을 가지고 있느냐에 따라 해당 공간의 차원이 결정된다.

구매 빈도

- (20, 30,000, 30)

- (23, 100,000, 21)

- (25, 74,000, 15)

구매 금액

- (60, 30,000, 3)   - (60, 300,000, 3)

나이

벡터와 관련된 용어들
벡터 공간에 대한 이해

- 벡터 공간(Vector Space) = 벡터들의 집합

- 크기(norm) = 벡터의 길이

- 방향(direction) = 단위 벡터(unit vector)

구매 빈도

- (20, 30,000, 30)

- (23, 100,000, 21)

- (25, 74,000, 15)

구매 금액

- (60, 30,000, 3)   - (60, 300,000, 3)

나이
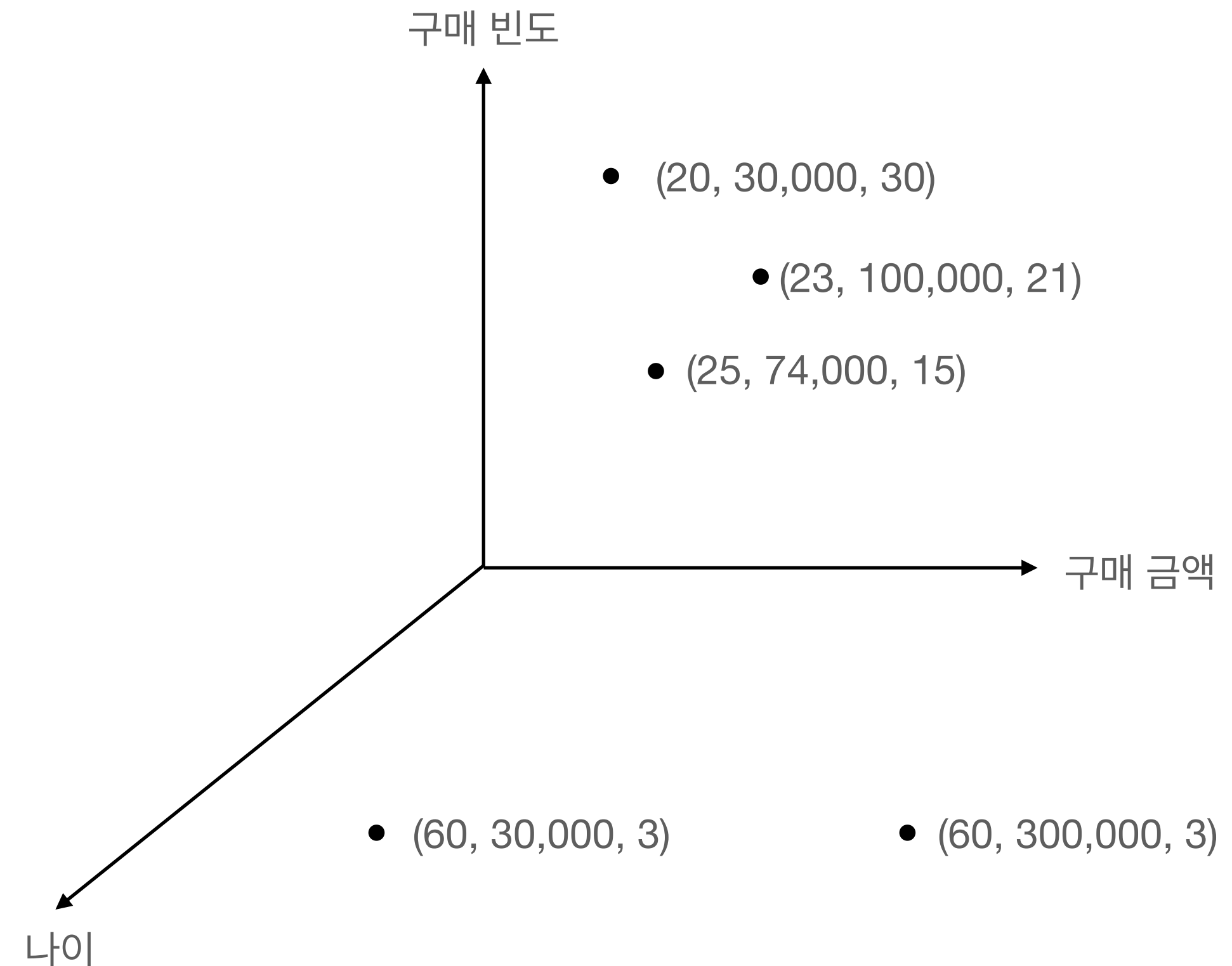
데이터 분석 예시

Q1. 다음 주어진 데이터를 벡터로 표현한다면 축은 무엇으로 잡을까?

Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

데이터 분석 예시

Q2. 다음 주어진 데이터는 몇 차원 벡터일까?

Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

데이터 분석 예시
Q3. 다음 데이터의 Calories column의 경우엔 공간에 어떻게 표현되는가?

Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

3. 데이터 분석을 위한 벡터 연산

벡터 기본 연산
데이터의 특징을 파악할 수 있는 기본 연산

- N차원의 벡터 $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)$와 $\boldsymbol{y} = (y_1, y_2, \cdots, y_N)$에 대해,

- 벡터의 크기 : $|\boldsymbol{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}$

- 벡터의 덧셈 : $\boldsymbol{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_N + y_N)$

- 벡터의 뺄셈 : $\boldsymbol{x} - \mathbf{y} = (x_1 - y_1, x_2 - y_2, \ldots, x_N - y_N)$

- 스칼라 배 : $a\boldsymbol{x} = (ax_1, ax_2, \cdots, ax_N)$

- 백터의 내적 : $x \cdot y = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$

데이터 분석 예시

Q4. Cheeseburger의 크기를 계산하여라.

Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

데이터 분석 예시

Q5. Cheeseburger와 Quarter Pounder 사이의 차이를 구하여라.

Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

데이터 분석 예시

Q6. 주어진 Cheeseburger, McChicken 데이터의 일부를 벡터로 표현할 때, 두 벡터의 내적값을 구하여라.

Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

벡터 간 거리
데이터 사이의 유사성 측정

- N차원의 벡터 $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)$와 $\boldsymbol{y} = (y_1, y_2, \cdots, y_N)$에 대해,

- Manhattan Distance (L1 distance) :

- Euclidean Distance (L2 distance) :

- Cosine Distance :

데이터 분석 예시

Q7. Cheeseburger와 Quarter Pounder 사이의 euclidean distance를 구하여라.

Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

데이터 분석 예시

Q7. Cheeseburger와 Quarter Pounder 사이의 cosine distance를 구하여라.

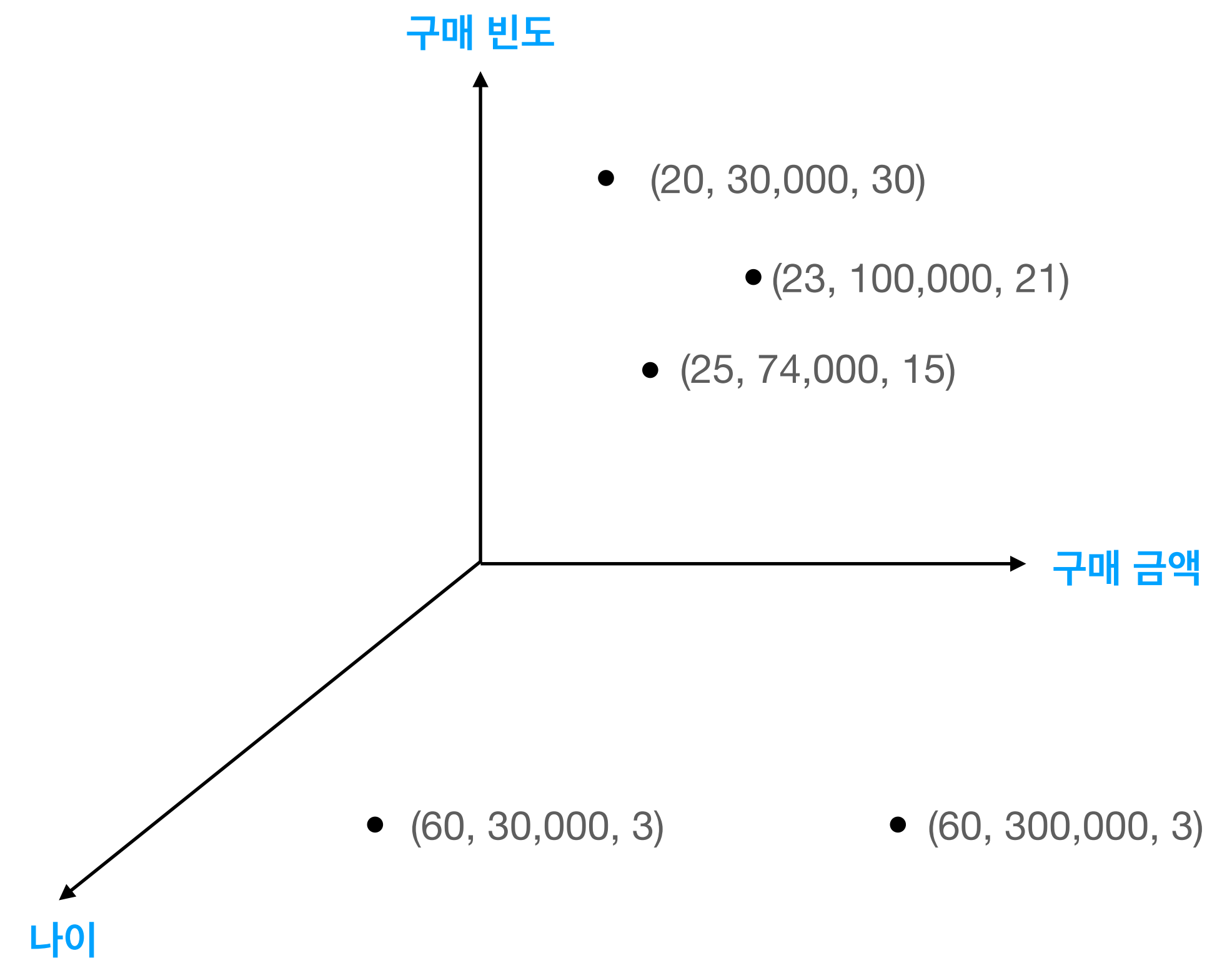| Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

4. Feature Space

Feature Space의 정의
주어진 데이터의 특징을 정의한 벡터 공간

- 주어진 데이터를 input vector라고 할 때, input vector 중에 필요한 특징만을 추출(또는 선별) 하여 벡터로 표현한 것을 "feature vector" 라고 한다.

- 필요한 특징을 선별하는 방법은 데이터를 잘 이해하고 있는 분석가가 담당한다.

- 필요한 특징을 추출하는 방법은 차원 축소 모델을 사용한다.

구매 빈도

- (20, 30,000, 30)
- (23, 100,000, 21)
- (25, 74,000, 15)

구매 금액

- (60, 30,000, 3)
- (60, 300,000, 3)

나이

데이터 분석 예시
column의 의미를 기준으로 선별하는 경우

Item	Calories	Calories from Fat	Total Fat (g)	Saturated Fat (g)	Trans Fat (g)	Cholesterol (mg)	Sodium (mg)	Carbs (g)	Fiber (g)	Sugars (g)	Protein (g)	Weight Watchers Pnts
Hamburger	250	80	9	3.5	0.5	25	520	31	2	6	12	247.5
Cheeseburger	300	110	12	6	0.5	40	750	33	2	6	15	297
Double Cheeseburger	440	210	23	11	1.5	80	1150	34	2	7	25	433
McDouble	390	170	19	8	1	65	920	33	2	7	22	383
Quarter Pounder® with Cheese	510	230	26	12	1.5	90	1190	40	3	9	29	502
Double Quarter Pounder® with Cheese	740	380	42	19	2.5	155	1380	40	3	9	48	720
Big Mac®	540	260	29	10	1.5	75	1040	45	3	9	25	534
Big N' Tasty®	460	220	24	8	1.5	70	720	37	3	8	24	452
Big N' Tasty® with Cheese	510	250	28	11	1.5	85	960	38	3	8	27	502
Angus Bacon & Cheese	790	350	39	17	2	145	2070	63	4	13	45	775
Angus Deluxe	750	350	39	16	2	135	1700	61	4	10	40	736
Angus Mushroom & Swiss	770	360	40	17	2	135	1170	59	4	8	44	751
Filet-O-Fish®	380	170	18	3.5	0	40	640	38	2	5	15	373.5
McChicken ®	360	150	16	3	0	35	830	40	2	5	14	354
McRib ®	500	240	26	10	0	70	980	44	3	11	22	499

데이터 분석 예시
추출 기법을 통하여 새로운 **Feature vector**를 생성한 경우 (e.g. PCA)

V1	V2	V3	V4	V5	V6	V7	V8
-1.3598071336738	-0.0727811733098497	2.53634673796914	1.37815522427443	-0.338320769942518	0.462387777762292	0.239598554061257	0.0986979012610507
1.19185711131486	0.26615071205963	0.16648011335321	0.448154078460911	0.0600176492822243	-0.0823608088155687	-0.0788029833323113	0.0851016549148104
-1.35835406159823	-1.34016307473609	1.77320934263119	0.379779593034328	-0.503198133318193	1.80049938079263	0.791460956450422	0.247675786588991
-0.966271711572087	-0.185226008082898	1.79299333957872	-0.863291275036453	-0.0103088796030823	1.24720316752486	0.23760893977178	0.377435874652262
-1.15823309349523	0.877736754848451	1.548717846511	0.403033933955121	-0.407193377311653	0.0959214624684256	0.592940745385545	-0.270532677192282
-0.425965884412454	0.960523044882985	1.14110934232219	-0.168252079760302	0.42098688077219	-0.0297275516639742	0.476200948720027	0.260314333074874
1.22965763450793	0.141003507049326	0.0453707735899449	1.20261273673594	0.191880988597645	0.272708122899098	-0.00515900288250983	0.0812129398830894
-0.644269442348146	1.41796354547385	1.0743803763556	-0.492199018495015	0.948934094764157	0.428118462833089	1.12063135838353	-3.80786423873589
-0.89428608220282	0.286157196276544	-0.113192212729871	-0.271526130088604	2.6695986595986	3.72181806112751	0.370145127676916	0.851084443200905
-0.33826175242575	1.11959337641566	1.04436655157316	-0.222187276738296	0.49936080649727	-0.24676110061991	0.651583206489972	0.0695385865186387
1.44904378114715	-1.17633882535966	0.913859832832795	-1.37566665499943	-1.97138316545323	-0.62915213889734	-1.4232356010359	0.0484558879088564
0.38497821518095	0.616109459176472	-0.874299702595052	-0.0940186259679115	2.92458437838817	3.31702716826156	0.470454671805879	0.53824722837695
1.249998742053	-1.21163680921816	0.383930151282291	-1.23489868766892	-1.48541947377961	-0.753230164566149	-0.689404975426345	-0.227487227519552
1.0693735878819	0.287722129331455	0.828612726634281	2.71252042961718	-0.178398016248009	0.337543730282968	-0.0967168617395962	0.115981735546597
-2.7918547659339	-0.327770756658658	1.64175016056605	1.76747274389883	-0.136588446465306	0.80759646826532	-0.422911389711497	-1.90710747624096
-0.752417042956605	0.345485415344747	2.05732291276727	-1.46864329840046	-1.1583936804082	-0.0778498291166733	-0.608581418236123	0.00360348436201849
1.10321543528383	-0.0402962145973447	1.2673320885949	1.28909146962552	-0.735997163604068	0.288069162976262	-0.586056786337461	0.189379713679593

데이터 분석 예시
학습을 통하여 새로운 Feature vector를 생성한 경우 (e.g. embedding)

feline friends say

Text

Embedding
model

| -0.001 | 0.006 | -0.014 | ... | -0.013 |

Text as vector

End of Slides